# Rajkumar Pujari

email: *rajkumarsaikorian@gmail.com* rajkumar-pujari.com Github: @pujari-rajkumar Phone: +1 (765)-701-7617

## EDUCATION

**Purdue University**, West Lafayette, Indiana, USA — Aug 2017–Present
Ph.D. in Natural Language Processing, Department of Computer Science — **3.80/4.0**

**Indian Institute of Technology Kharagpur**, West Bengal, India — Jul 2010–May 2014
B.Tech. (Hons.), Department of Computer Science and Engineering — **9.10/10**

## PUBLICATIONS

**LLM-Human Schema Pipeline for Cultural Context Grounding of Conversations**
Rajkumar Pujari and Dan Goldwasser
Accepted to NAACL 2025 main conference, Albuquerque, USA April 29–May 4

***"We Demand Justice!"*: Towards Social Context Grounding of Political Texts**
Rajkumar Pujari, Chengfei Wu and Dan Goldwasser
Talk at EMNLP 2024 main conference, Miami, USA November 12–16

**Reinforcement Learning Guided Multi-Task Learning for Low-Resource Stereotype Detection**
Rajkumar Pujari, Erik Oveson, Priyanka Kulkarni and Elnaz Nouri
Poster at ACL 2022 main conference, Dublin, Ireland from May 22–27

**Understanding Politics via Contextualized Discourse Processing**
Rajkumar Pujari and Dan Goldwasser
Talk at EMNLP 2021 main conference, Punta Cana, Dominican Republic on November 6–12

**Using Natural Language Relations between Answer Choices for Machine Comprehension**
Rajkumar Pujari and Dan Goldwasser
Talk at NAACL-HLT 2019 main conference, Minneapolis, USA on June 2–7

**Can Taxonomy help? Improving Semantic Question Matching using Question Taxonomy**
Deepak Gupta, Rajkumar Pujari, Asif Ekbal, Pushpak Bhattacharyya, Anutosh Maitra, Tom Jain and Shubhashis Sengupta
Poster at COLING 2018 main conference, Santa Fe, New Mexico, USA, August 20–24

**A Novel Two-stage Framework for Extracting Opinionated Sentences from News Articles**
Rajkumar Pujari, Swara Desai, Niloy Ganguly and Pawan Goyal
Talk at Texgraphs-9 workshop at EMNLP 2014, Doha, Qatar, October 25–29

## RESEARCH EXPERIENCE

**Ph.D. Student, Purdue University** — Aug 2017–Present
*Contextualized Discourse Representations; Machine Comprehension; Social Norm Understanding (Pragmatics )*
Prof. Dan Goldwasser
Worked on investigating the usefulness of common sense knowledge and inferences in machine comprehension. Currently working on developing a distributed, interpretable representation for large political corpora and also, working on a DARPA CCU (Computational Cultural Understanding) project that leverages Large Language Models combined with symbolic models to extract social norms and their violations from conversations.

**Project Research Assistant, CFILT, IIT Bombay** — Jan 2016–May 2017
*Semantic Question matching using a taxonomy and DL representations* — Prof. Pushpak Bhattacharyya
Developed a restricted-domain QA system for Accenture labs, Bangalore. Was primarily focused on developing a hierarchical taxonomy and subsequent classification algorithms used for semantic question matching.

**Bachelor Thesis Project, IIT Kharagpur** — Jul 2013–May 2014
*Extracting opinionated sentences from news articles* — Prof. Niloy Ganguly
Developed a novel two-stage framework to extract opinion-bearing sentences from news articles. The framework is a pipeline of sentiment classifier and a graph-based algorithm analogous to HITS in collaboration with Yahoo!.

## RELEVANT INTERNSHIPS

**JEM Intern, Microsoft Research, Redmond** — May 2021–Aug 2021
*RL-guided Multitask model for Stereotype Detection in Text* — Elnaz Nouri, Erik Oveson and Priyanka Kulkarni
Worked on Reinforcement Learning guided Multitask Learning Model for *Stereotype Detection*. Leveraged several toxicity datasets such as *Hate-Speech, Misogyny & Offensive Language* to systematically improve performance on the task of *Stereotype Detection*. Built a *Stereotype Detection* dataset using Amazon Mechanical Turk.

**Applied Scientist Intern, Amazon Alexa Coversational Search Team** — May 2019–Aug 2019
*Conversational Question Answering in Sports Domain* — Kevin Small
Developed a new conversational QA dataset for sports domain. Built a ParlAI-style data collection tool and collected the dataset using Amazon MTurk. Also designed a generative neural architecture for the task.

**Summer Internship, Yahoo! Bangalore, India** May 2013–Jul 2013
*Relevance ranking of comments on news articles* Ms.Swara Desai
Designed and implemented a comment ranking algorithm that scores comments based on the relevance of the comment text to the article and its Yahoo classification category (YCT).

## Positions of Responsibility

**Program Committee Member - Reviewer; *Outstanding Reviewer**
ACL ARR since Nov 2021, ACL (2024, 2023, 2022), ACL-IJCNLP 2021*, EMNLP (2024, 2023, 2022, 2021, 2020), EMNLP-IJCNLP 2019, NAACL-HLT (2021, 2019), NAACL 2022, AAAI (2023, 2022, 2021, 2020), COLING (2025, 2022, 2020), LREC-COLING 2024, IJCAI (2024, 2023, 2020), TALLIP

**Graduate Teaching Assistant** Jul 2017–Present
Data Mining and Machine Learning, Web Information Search and management, Data structures and algorithms, and Operating Systems undergraduate courses.

**President, Purdue University Cricket Club** May 2019–Jun 2021
Organized cricket tournaments for 250 players, raised funds, and managed an annual budget of USD $7000.

## Work Experience

**Senior Quantitative Researcher** Jul 2014–Dec 2015
WorldQuant Research, India
Was responsible for researching financial and mathematical literature, understanding various datasets to identify sources of market inefficiencies and convert them to predictive profitable models called *alpha*s. The objective was to identify and construct signals and make robust models from them with high Sharpe ratios (returns/risk) and significant abnormal returns. Concentrated mainly on seeking low turnover quality *alphas* for trading in the equity market which are used in developing algorithmic daily re-balancing long-short trading strategies on US, Europe, Asian, and other markets

## Notable Achievements

Was promoted to Senior Quantitative Researcher after $1^{st}$ year at Worldquant Research for exceptional performance

Secured a Department Change from Electrical Engineering department to Computer Science department at the end of the first year based on academic merit ($5\%$ acceptance rate)

Secured $840^{th}$ rank amongst more than 450,000 students (99.81 percentile) in IIT-JEE 2010 and $939^{th}$ rank amongst more than 1,000,000 students (99.91 percentile) in AIEEE 2010

Qualified among Center Top 10% in National Standard Examination in Physics and National Standard Examination in Chemistry conducted by Indian Association of Physics Teachers (IAPT) in Class XII standard.

## Key Courses

**Graduate Level:** Advanced Topics In Reasoning with LLMs (A+), ML Methods for NLP, Statistical ML (A+), Deep Learning, Numerical Methods for Optimization, Algorithm Design and Analysis, Operating Systems

**Data Mining:** Information Retrieval, Machine Learning, Speech and Natural Language Processing

**Electives:** Advanced Graph Theory, Artificial Intelligence, Computational Number Theory, Database Management and Systems, Distributed Systems, Formal Systems, Foundations of Cryptography

## Key Projects

**Summer Project, IIT Kharagpur** May 2012–Jun 2012
*Tracking soccer players in a video* Prof.Partha Pratim Das
Worked on various color segmentation algorithms and implemented a tracking algorithm based on position prediction in successive frames using MATLAB Image Processing Toolkit

## Extra-Curricular Activities

Represented Purdue Cricket Club in MWCC (2018) and MCT (2019, 2021).

Senior Writing Team Member of Entrepreneurship Cell IIT Kharagpur in the academic year 2011–2012

National Service Scheme (NSS) volunteer in the Education Improvement Group and a National Cadet Corps (NCC) cadet for 6 years and was awarded NCC-**A** & **B** certificates

Part of a 6-member team that qualified for the finals of the NIGHTSHIFT event in Kshitij 2012, the annual techno-management fest of IIT Kharagpur

## References

Prof. Dan Goldwasser, Purdue University email: *dgoldwas@purdue.edu*
Leora Morgenstern, SRI email: *leora.morgenstern@sri.com*
Elnaz Nouri, Microsoft Research email: *elnaz.nouri@microsoft.com*